

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/120238>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.



When the blind curve is finite: dimension estimation and model inference based on empirical waveforms

Fred Hasselman*

Learning and Plasticity, Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands

*Correspondence: f.hasselmann@bsi.ru.nl

Edited by:

Michael A. Riley, University of Cincinnati, USA

Reviewed by:

Michael A. Riley, University of Cincinnati, USA

John G. Holden, University of Cincinnati, USA

Sebastian Wallot, Aarhus University, Denmark

A PLEA FOR STRONG INFERENCE

The recent research topic “*Fractal Analyses: Statistical and Methodological Innovations and Best Practices*” reveals there is no consensus among experts about the best procedure to estimate self-affine structure in trial and time series data. One of the recurring issues pertains to the validity of inferences based on analysis results about the physical change processes that generated the empirical waveforms. In this paper I argue that none of these approaches can be used to validate such inferences outside of the context of theory evaluation by strong inference (e.g., Platt, 1964). Two arguments warrant this claim: (1) All procedures make an assumption about the physics of the system under scrutiny. This is arguably most prominent in ARFIMA modeling, but associating an estimated scaling exponent to a fractal dimension is also based on assumptions (e.g., fGn vs. fBm; Mandelbrot and Van Ness, 1968); (2) given infinitesimal measurement resolution and infinite observation time, properties like dimension and self-affinity are not unique descriptors of a process, pattern or object (cf. Vicsek, 2001). Multiple mathematical models of physical processes can be constructed to generate a waveform with exactly the dynamical and invariant properties as observed in the finite sample (e.g., Mandelbrot, 2001; Kantz and Schreiber, 2003; Thornton and Gilden, 2005; Morrison, 2008).

The second issue pertains to a general problem of model-based inference: a good fit to a finite sample of measurement outcomes can never be conclusive in the evaluation of predictions by theories (cf. Roberts and Pashler, 2000, 2002; Fiedler et al., 2012). Using results

of (fractal) analyses to answer questions about the physics of the observed system is an attempt to evaluate the ontology of a theory, ex post facto; let's leave ontology evaluation to the metaphysicians (cf. Poincaré, 1905, p. 211). The scientific method is not a competition for *mathematical models* constructed to produce the best fit to measurement outcomes; instead, *theoretical predictions* about the observed system compete for highest empirical precision and accuracy in order to gain scientific credibility.

In what follows I evaluate to what extent fractal analyses are used in the context of strong inference given the current empirical record of human physiology and performance. Subsequently I will explore what may be gained when implicit ontology falsification is removed from fractal analyses by introducing the concepts of intuitive dimension and informed dimension estimates.

ON FRACTAL SCALING AND PLANETARY ORBITS

Why should an accurate prediction by a theory be preferred over a good retrospective model fit? Models proposed to explain the orbit of Mercury (which displays a perihelion advance) present an interesting historical analogy. The orbit was accurately modeled by the classical geocentric models based on Ptolemy's *Almagest* (used from around 100–3500 CE; Toomer, 1984). These models assumed celestial objects moved around the earth on a celestial sphere that could host one or more local orbits or epicycles. The number of nested epicycles was simply varied until the predicted trajectory was sufficiently in accordance with

the empirical record. Curiously, to the heliocentric models replacing Ptolemaic astronomy like Newton's theory of celestial mechanics, Mercury's orbit was an anomaly! No wonder that Einstein considered the accurate prediction of this anomaly the most important empirical test of his theory of general relativity (Einstein, 1916; Will, 2005).

This brief history of orbit modeling reveals that the theoretical perspective used to observe the empirical record changes one and the same reliably measured pattern from a good model fit, into an anomalous phenomenon into a critical benchmark for theory evaluation. Ptolemy's solution of adding epicycles to reconstruct the shape of a trajectory is essentially the same as adding weighted autoregressive, moving average, seasonally changing, or (fractionally) integrating components in a time series model. Those components are constructed into the model in order to create a better fit with a pattern in the data. This is allowed by mathematics, but their presence is not predicted by a theory of principles about physical change processes in living systems and this renders its scientific evaluation invulnerable to the presence of anomalies. Compare to Newton's closed theory of principles: “*In Ptolemy's case, if the orbit didn't fit, he could add other epicycles. But if an experiment does not fit in Newtonian physics, you don't know what you mean by the words.*” (Heisenberg interviewed by Kuhn, 1963, p. 24, February 27th).

In order to advance scientific knowledge about scaling phenomena in living systems a program of strong inference that aims to produce closed theories of principles is needed. In order to reach

this goal, empirical inquiries need to go beyond describing scaling phenomena in different populations in the context of impaired performance or pathology (e.g., Goldberger et al., 2002; Gilden and Hancock, 2007; West, 2010; Wijnants et al., 2012a). Several recent studies reveal scaling phenomena can be brought under experimental control, which is essential for a program of strong inference (e.g., Kello et al., 2007; Wijnants et al., 2009; Van Orden et al., 2010; Correll, 2011; Holden et al., 2011; Kuznetsov et al., 2011; Stephen et al., 2012). The diverging theoretical predictions examined in most studies reveal that the observed waveforms are more likely to originate from interaction-dominant complexity than from component-dominant mechanics (also see Turvey, 2007; Kello et al., 2010; Diniz et al., 2011).

A closed theory should account for most phenomena in the existing empirical record. A first step was recently made in which it was shown that the well-known speed-accuracy tradeoff in human performance is meaningfully related to the emergence of self-affine structure via nested timescales (Wijnants et al., 2012b). At the current level of scientific understanding it seems reasonable to ask of those who insist models based on AR-processes provide a parsimonious explanation of fractal scaling (e.g., Wagenmakers et al., 2005; Torre and Wagenmakers, 2009; Stadnitski, 2012), to provide experimental evidence that can validate their claims.

As stated above however, most claims about scaling phenomena based on fractal analyses are prone to implicit ontology falsification. In what follows, I will suggest an approach to dimension estimation that is based on intuitions about the geometry of a curve rather than on known mathematical models of change processes. I will focus on the mono-fractal case and show that a consistent conversion scheme for common estimates of self-affine structure is possible when using this notion of dimension.

FINITE SELF-AFFINITY: THE BLIND CURVE AND THE PERIMETER WALK

Dimension is an intrinsic property of a mathematical object that indicates to what extent it occupies the topological space in which it is embedded. A dimension estimate that is based on the properties of an

empirical waveform can be defined as a finite walk in the plane that never forms a perimeter. Formally, this is a self-avoiding open curve dividing a bounded plane in two unconnected regions (i.e., it is not a Jordan Curve). Note that the properties of the curve have a physical origin: it is self-avoiding and open due to the arrow of time and because observation duration and measurement outcomes are finite, the planar topology is bounded.

Estimation procedures derived from formal definitions of dimension respect an intuitive geometric notion of a scaling of bulk with size (Theiler, 1990). Using the definition above, the intuitive concept to quantify would be a characterization of the waveform as line-like or plane-like, hence *planar extent* (e.g., Higuchi, 1988; Katz, 1988; Raghavendra and Dutt, 2010). Sevcik (1998) introduced such a dimension estimate based on the Hausdorff-Besicovitch dimension (Hausdorff, 1919; Xiao, 2008). It involves a double linear transformation of the axes embedding the waveform in a unit square of size N by N . Its length can be calculated as the sum of the Euclidean distance between points on the normalized curve. The graph entitled “Sevcik method” in **Figure 1** shows the equation used to approximate D based on number of observation intervals ($N-1$) and curve length L (for details see Sevcik, 1998, 2006). Across the top of **Figure 1**, twelve different waveforms are shown that were analyzed for self-affine structure (see caption for details). The waveforms were generated using freely available Matlab scripts¹.

INTUITIVE DIMENSION

The dimension estimates for the 12 waveforms based on Sevcik’s method are numerically different from D_{uBC} , as well as from known exact values, the goal however is to achieve relative consistency. Processes that generated waveform 9 and 10 have known $D = 1.5$, which is equal to a sequence of random numbers drawn from a Normal distribution (i.e., waveform 5, waveform 10 is its cumulative sum). Using Sevcik’s method however, both waveforms are classified as a Brownian noise. If this waveform were known to be

physiological and medical in nature, the constrained dynamics associated with Brownian noise would lead to profoundly different conclusions about the health and well-being of the patient in question compared to blindly interpreting the limit values $D = 1.5$ and $H = 0.5$ (e.g., Goldberger et al., 2002; Van Orden et al., 2009). The gray-scale areas represent Sevcik’s estimate for 12100 simulated series with ideal spectral slopes ranging from -3 to 3 . Note that at 25% of the data length (first set of markers) the relative ordering according to D_{uBC} is recovered for almost all waveforms.

INFORMED DIMENSION ESTIMATES

The other graphs represent self-affinity exponents estimated using the power spectrum (PSD), detrended fluctuation analysis (DFA), standardized dispersion analysis (SDA), and ARFIMA (modeling strategy: Reisen and Lopes, 1999; Silva et al., 2006). The informed estimates of D refer to conversions of the self-affinity indices obtained for the 12 waveforms. ARFIMA modeling did not provide a consistent conversion scheme therefore the differencing parameter was plotted against the PSD based estimates. The gray-scale coded regions in these plots refer to the PSD slope estimates for the 12100 simulated series. These areas thus display the relation between the DFA, SDA, and ARFIMA self-affinity indices and the PSD slope based estimate.

The equations suggested for PSD and DFA indices produce approximations of well-known (H , D) pairs and should not be confused with an analytic solution. For SDA the known formula *1-Slope* yields relatively consistent results, a problem is that for some ranges of PSD slopes, the SDA indices are the same. ARFIMA modeling by the AIC selection criterion preferred models without self-affine structure ($d = 0$) of varying order ($p = 0-2$, $q = 0-2$) for the majority of simulated PSD slope series (ARIMA 42.8% vs. ARFIMA 20.6%). The remaining series produced fit errors (see Supplementary Materials for details).

Most waveforms in **Figure 1** get assigned a value for D that is in accordance with their planar extent as indicated by D_{uBC} . Waveform 11 is more line-like than waveform 9 and 10, which both map

¹The scripts are available at <http://fredhasselmann.com> in the section Supplementary Materials.

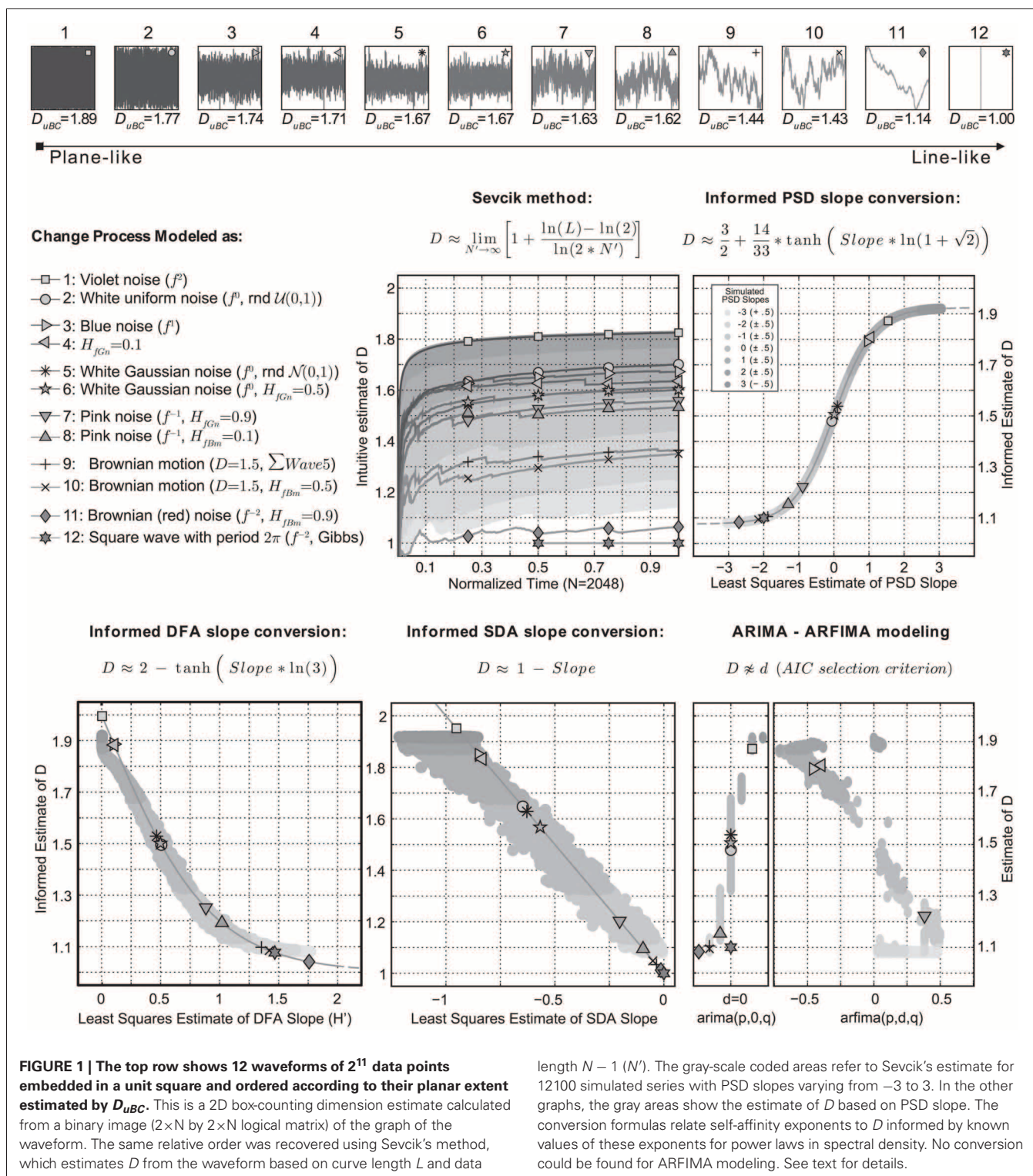


FIGURE 1 | The top row shows 12 waveforms of 2^{11} data points embedded in a unit square and ordered according to their planar extent estimated by D_{uBC} . This is a 2D box-counting dimension estimate calculated from a binary image ($2 \times N$ by $2 \times N$ logical matrix) of the graph of the waveform. The same relative order was recovered using Sevcik's method, which estimates D from the waveform based on curve length L and data

length $N - 1$ (N'). The gray-scale coded areas refer to Sevcik's estimate for 12100 simulated series with PSD slopes varying from -3 to 3 . In the other graphs, the gray areas show the estimate of D based on PSD slope. The conversion formulas relate self-affinity exponents to D informed by known values of these exponents for power laws in spectral density. No conversion could be found for ARFIMA modeling. See text for details.

closely to known D of Brownian noise. Expected exceptions are waveforms 2 and 12. A sequence of random numbers drawn from a uniform distribution (waveform 2)

has a PSD slope of zero. Taking the Fourier transform of a square wave (waveform 12) gives a frequency spectrum of odd harmonics only, with a slope of exactly

-2 (the Gibbs phenomenon). Another expected result is that ARFIMA is preferred for series produced as H_{fGn} (except $H_{fGn} = 0.5$).

TO SELF-AFFINITY ... AND BEYOND!

It seems possible to remove implicit assumptions about system ontology from fractal analysis by defining dimension as the planar extent of a finite curve. Direct estimates based on curve length and 2D box-counting provide a consistent relative ordering on this dimension. An informed conversion scheme using estimates of self-affine structure obtained from PSD, DFA, and SDA analyses give similar results. Some exceptions were predicted, but ARFIMA modeling could not be included in the approach due to inconsistent analysis results. A mono-fractal perspective was explored here, but there is no reason to assume it cannot be extended to the multi-fractal framework as well.

Exact numerical similarity of estimates is sacrificed for the convergence of estimates to a similar relative ordering. This sacrifice is acceptable given that in principle, even the best estimates of dimension and self-affinity leave us blind to the physical processes that generated the waveform. I suggest that claims about the physics of the system need to be evaluated by comparing the empirical accuracy of theoretical predictions in a program of strong inference, not by comparing fit indices.

REFERENCES

- Correll, J. (2011). Order from chaos? 1/f Noise predicts performance on reaction time measures. *J. Exp. Soc. Psychol.* 47, 1–6.
- Diniz, A., Wijnants, M. L., Torre, K., Barreiros, J., Crato, N., Bosman, A. M. T., et al. (2011). Contemporary theories of 1/f noise in motor control. *Hum. Mov. Sci.* 30, 889–905.
- Einstein, A. (1916). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik.* 354, 769–822.
- Fiedler, K., Kutzner, F., and Krueger, J. I. (2012). The long way from α -error control to validity proper: problems with a short-sighted false-positive debate. *Perspect. Psychol. Sci.* 7, 661–669.
- Gilden, D. L., and Hancock, H. (2007). Response variability in attention-deficit disorders. *Psychol. Sci.* 18, 796–802.
- Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., and Stanley, H. E. (2002). Fractal dynamics in physiology: alterations with disease and aging. *Proc. Natl. Acad. Sci. U.S.A.* 99(Suppl. 1), 2466–2472.
- Hausdorff, F. (1919). Dimension and äusseres Mass. *Math. Ann.* 79, 157–179.
- Heisenberg, W. (1963). *Oral History Interview of Werner Heisenberg by Thomas Kuhn. Archive for the History of Quantum Physics.* Cambridge, MA: Harvard University. [deposit at Harvard University, Cambridge, MA].
- Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D* 31, 277–283.
- Holden, J. G., Choi, I., Amazeen, P. G., and Van Orden, G. (2011). Fractal 1/f dynamics suggest entanglement of measurement and human performance. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 935–948.
- Kantz, H., and Schreiber, T. (2003). *Nonlinear Time Series Analysis, 2nd Edn.* Cambridge, UK: Cambridge University Press.
- Katz, M. J. (1988). Fractals and the analysis of waveforms. *Comput. Biol. Med.* 18, 145–156.
- Kello, C. T., Beltz, B. C., Holden, J. G., and Van Orden, G. C. (2007). The emergent coordination of cognitive function. *J. Exp. Psychol. Gen.* 136, 551–568.
- Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., et al. (2010). Scaling laws in cognitive sciences. *Trends Cogn. Sci.* 14, 223–232.
- Kuznetsov, N. A., Wallot, S., Alexander, E., and Ihlen, F. (2011). Effects of accuracy feedback on fractal characteristics of time estimation. *Front. Integr. Neurosci.* 5:62. doi: 10.3389/fnint.2011.00062
- Mandelbrot, B. B. (2001). Scaling in financial prices: III. Cartoon Brownian motions in multifractal time. *Quant. Finance* 1, 427–440.
- Mandelbrot, B. B., and Van Ness, J. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Rev.* 10, 422–437.
- Morrison, F. (2008). *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness and Determinism.* Mineola, NY: Dover Publications, Inc.
- Platt, J. (1964). Strong inference. *Science* 146, 347–353.
- Poincaré, H. (1905). *Science and Hypothesis (Trans. p. Science and Hypothesis).* New York, NY: The Walter Scott Publishing Co., Ltd.
- Raghavendra, B. S., and Dutt, D. N. (2010). Computing fractal dimension of signals using multiresolution box-counting method. *Int. J. Inf. Math. Sci.* 6, 50–65.
- Reisen, V. A., and Lopes, S. (1999). Some simulations and applications of forecasting long-memory time-series models. *J. Stat. Plan. Inference* 80, 269–287.
- Roberts, S., and Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychol. Rev.* 107, 358–367.
- Roberts, S., and Pashler, H. (2002). Reply to Roders and Rowe (2002). *Psychol. Rev.* 109, 605–605.
- Sevcik, C. (1998). A procedure to estimate the fractal dimension of waveforms. *Complexity International*, 05. Available online at: <http://www.complexity.org.au/ci/vol05/sevcik/sevcik.html>
- Sevcik, C. (2006). On fractal dimension of waveforms. *Chaos Solitons Fractals* 28, 579–580.
- Silva, E. M., Franco, G. C., Reisen, V. A., and Cruz, F. R. B. (2006). Local bootstrap approaches for fractional differential parameter estimation in ARFIMA models. *Comput. Stat. Data Anal.* 51, 1002–1011.
- Stadnitski, T. (2012). Measuring fractality. *Front. Physiol.* 3:127. doi: 10.3389/fphys.2012.00127
- Stephen, D. G., Anastas, J. R., and Dixon, J. A. (2012). Scaling in cognitive performance reflects multiplicative multifractal cascade dynamics. *Front. Physiol.* 3:102. doi: 10.3389/fphys.2012.00102
- Theiler, J. (1990). Estimating fractal dimension. *J. Opt. Soc. Am. A* 7, 1055.
- Thornton, T. L., and Gilden, D. L. (2005). Provenance of correlations in psychological data. *Psychon. Bull. Rev.* 12, 409–441.
- Toomer, G. J. (1984). *Ptolemy's Almagest.* New York, NY: Springer-Verlag.
- Torre, K., and Wagenmakers, E.-J. (2009). Theories and models for 1/f(beta) noise in human movement science. *Hum. Mov. Sci.* 28, 297–318.
- Turvey, M. T. (2007). Action and perception at the level of synergies. *Hum. Mov. Sci.* 26, 657–697.
- Van Orden, G. C., Kello, C. T., and Holden, J. G. (2010). Situated behavior and the place of measurement in psychological theory. *Ecol. Psychol.* 22, 24–43.
- Van Orden, G. C., Kloos, H., and Wallot, S. (2009). “Living in the pink: intentionality, wellbeing, and complexity,” in *Handbook of the Philosophy of Science*, ed C. Hooker (Amsterdam: Elsevier), 639–682.
- Vicsek, T. (2001). *Fluctuations and Scaling in Biology.* Oxford: Oxford University Press.
- Wagenmakers, E.-J., Farrell, S., and Ratcliff, R. (2005). Human cognition and a pile of sand: a discussion on serial correlations and self-organized criticality. *J. Exp. Psychol. Gen.* 134, 108–116.
- West, B. J. (2010). Homeostasis and Gauss statistics: barriers to understanding natural variability. *J. Eval. Clin. Pract.* 16, 403–408.
- Wijnants, M. L., Bosman, A. M. T., Hasselman, F., Cox, R. F. A., and Van Orden, G. C. (2009). 1/f scaling in movement time changes with practice in precision aiming. *Nonlin. Dyn. Psychol. Life Sci.* 13, 79–98.
- Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., and Van Orden, G. (2012a). An interaction-dominant perspective on reading fluency and dyslexia. *Ann. Dyslexia* 62, 100–119.
- Wijnants, M. L., Cox, R. F. A., Hasselman, F., Bosman, A. M. T., and Van Orden, G. (2012b). A trade-off study revealing nested timescales of constraint. *Front. Physiol.* 3:116. doi: 10.3389/fphys.2012.00116
- Will, C. (2005). “Special relativity: a centenary perspective,” in *Einstein, 1905–2005: Poincaré Seminar 2005*, Vol. 47, eds T. Damour, O. Darrigol, B. Duplantier, and V. Rivasseau (Basel-Boston-Berlin: Birkhäuser Verlag), 33–58.
- Xiao, Y. (2008). Packing dimension, Hausdorff dimension and Cartesian product sets. *Math. Proc. Camb. Philos. Soc.* 120, 535.

Received: 18 January 2013; accepted: 21 March 2013;
published online: 08 April 2013.

Citation: Hasselman F (2013) When the blind curve is finite: dimension estimation and model inference based on empirical waveforms. *Front. Physiol.* 4:75. doi: 10.3389/fphys.2013.00075

This article was submitted to *Frontiers in Fractal Physiology*, a specialty of *Frontiers in Physiology*. Copyright © 2013 Hasselman. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.